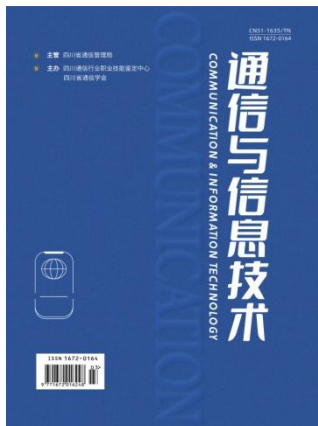




通信与信息技术

Communication & Information Technology

国内统一连续出版物号：CN 51-1635/TN

国际标准出版物号：ISSN 1672-0164

邮发代号：62-166

题目：基于多维度协同策略的气象HPC全栈式算力优化方法研究

作者：王晓东，邓卫华，李洪康，熊虹华

优先出版日期：2025年5月28日

万方网站：<https://w.wanfangdata.com.cn/>

优先出版：优先出版是指编辑部录用并定稿的文章，通过具备网络出版资质的数字出版平台，先于印刷版杂志出版日期出版，文章内容、排版已定稿，视作正式出版。为确保录用定稿优先出版文章的严肃性，文章一经发布，不得修改题目、作者、作者排序、工作单位，只可基于编辑规范进行少量文字修改。

《通信与信息技术》为双月刊，逢单月底出刊，是国内外公开出版的自然科学学术期刊，设置了运营一线、热点技术、行业观察、解决方案、专网通信等栏目。

办刊宗旨：面向行业，沟通社会；宣传政策，促进发展；为通信发展服务，为通信企业服务，为通信科技人员和职工服务，为广大通信消费者服务，集信息性、行业性、技术性为一体的综合类通信刊物。

基于多维度协同策略的气象 HPC 全栈式算力优化方法研究

王晓东, 邓卫华, 李洪康, 熊虹华

江西省气象数据中心, 江西南昌 330096

摘要: 随着机器学习与大数据技术的快速发展, 气象领域对算力资源的需求呈现指数级增长。在硬件资源受限的条件下, 如何提升高性能计算集群(HPC)的算力利用效率成为关键研究课题。基于某省级气象 HPC 集群的体系架构与业务特征, 提出多维度协同优化框架, 从硬件资源调度、系统配置调优、动态任务管理及编译效率提升四个核心层面展开研究。通过构建层次化优化模型, 实现了存储 I/O 吞吐性能、作业调度效率、编译执行效率显著提升的效果。本研究成果为气象领域 HPC 系统的效能优化提供了可复用的技术路径。

关键词: 高性能计算; 算力优化; 气象应用; 异构资源调度

中图分类号: TP393

文献标志码: A

1 引言

高性能计算集群 (High Performance Computing Cluster, HPC) 是指由多台高性能计算机、网络系统、存储系统组成的计算集群, 采用先进并行计算模式和作业调度策略搭建完成大规模科学计算环境, 用于处理复杂计算问题。作为支撑现代气象数值预报的核心基础设施, HPC 算力效能直接影响数值天气预报 (NWP) 的时空分辨率与计算时效性。当前气象 HPC 系统面临三大挑战: 一是多源异构气象数据的爆炸式增长导致存储 I/O 瓶颈; 二是实时业务与科研计算任务的资源竞争; 三是国产化硬件平台与气象数值模式的适配优化需求。

王洁^[1]、张云泉^[2]等、高岳^[3]等、杨敏^[4]等, 从国内高性能计算机发展现状、国产化发展策略、GPU 集群应用, 以及 HPC 与大数据、AI 等新型技术融合情况进行研究。王亚娟^[5]、张宇^[6]等系统论证了 HPC 在气象海量资料处理、气象灾害预警等方面应用。明宇^[7]、张志坚^[8]等、徐顺^[9]等深入分析了高性能异构计算调度算法、软件, 以及异构计算架构在气象领域的适用性。杨昕^[10]等、贾晨刚^[11]等、沈瑜^[12]等、顾文静^[13]等对国家级、省级气象 HPC 建设情况, 从集群架构、系统扩容、统一资源管理等角度对气象应用进行研究。黄小猛^[14]等、雷鸣^[15]等对 AI+ 气象模型发展应用进行了研究, 展示了人工智能时代气象 HPC 承载业务的新需求。李江南^[16]、吴

璨^[17]等均对 HPC 作业调度方法开展了研究, 其中李江南^[16]提出一种自适应调度算法, 提升作业调度性能, 吴璨等^[17]提出一种面向高性能计算环境的智能任务编排架构, 可以自动解析作业之间的依赖关系, 智能编排作业提交顺序, 具备较好应用效果。田彦明等^[18]从计算节点负载均衡角度出发, 对均衡分配算法进行深入研究。吕小敬等^[19]从集群应用的角度, 从动力过程优化、物理过程优化、支撑技术优化等几个方面, 对区域气候模式 (CWRF) 并行优化方案进行详细研究。

然而, 现有研究多聚焦单一优化维度, 缺乏针对气象业务特征的协同优化方案。本研究创新提出“四维协同”优化框架, 在省级气象 HPC 集群中实现全栈效能提升。

2 气象 HPC 系统特性与架构分析

2.1 气象业务特征

省级气象 HPC 系统呈现三大典型特征:

一是任务类别多样性。某省气象运行业务包含省级短临快速更新循环同化预报系统 (WRF-REFRESH)、省级中尺度数值预报系统、省级多源数据融合系统等实时气象业务, 同时还有其他各类服务、决策和科研业务。不同任务对异构资源 (CPU、内存、存储、网络) 需求和优先级差异显著。

二是气象数据密集型。计算作业大多需要包含地面观测、雷达、高空、卫星等海量气象数据的读写和处理, 对 I/O 存

收稿日期: 2025 年 3 月 17 日; 修回日期: 2025 年 5 月 12 日

通信作者: 邓卫华 (1977—), 女, 硕士, 高级工程师, 研究方向: 气象信息技术研究。

储能力和GPU处理所属访存要求较高,如雷达基数据与卫星遥感数据日处理量可达TB级,存储持续吞吐量要求 $\geq 50\text{GB/s}$ 。

三是计算时效敏感性。上述计算作业,既有固定时段(如每天、每小时固定时间)运行的实时任务,如短临预报需满足10分钟级计算周期,台风路径模拟需保障100节点级并行效率,同时也有用于科研的作业需在不定期短时间大量占用计算资源,还有突发性灾害预警任务需实现计算资源的抢占式调度。

2.2 气象 HPC 体系架构

本文实验环境以某省级气象HPC为例,该集群于2014年完成建设投入业务使用,集群包含42台“刀片式”计算节点+1台“胖”计算节点构成,网络系统使用1套56Gbps FDR

Infiniband网络作为数据通信网络,1套千兆以太网网络用于系统管理,系统监控软件使用曙光Gridview集群管理监控软件、IBM Platform PAC网页监控软件。集群架构及网络拓扑见图1,具体硬件参数包括:

计算层: 42台曙光CB50-G20双路刀片计算节点(2.5GHz/64GB)、1台Power 822L双路服务器(3.42GHz/128GB)。

存储层: 双活架构曙光DS800-G25存储系统,2台曙光I620-G20双路IO节点,202TB裸容量,支持Lustre并行文件系统。

网络层: 1套56Gb FDR InfiniBand互连提升MPI程序执行效率,1套千兆以太网网络用于系统管理。

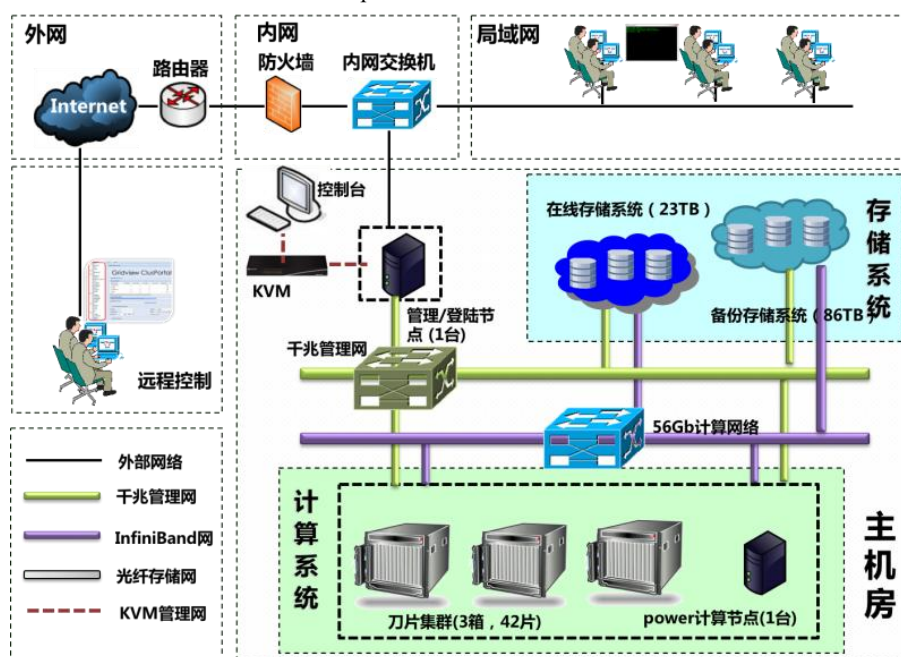


图1 某省气象高性能计算机集群架构图

本文针对气象HPC架构特性,从硬件资源、作业调度、编译层、代码优化等多种维度研究分析优化方法。

3 集群硬件层优化

硬件资源是HPC集群性能的物理基础,在硬件层面优化可大致分为计算节点BIOS优化、存储优化、GPU优化、网络优化等。

3.1 计算节点 BIOS 设置优化

节点服务器默认硬件工作模式均为均衡或节能模式,为提升工作效能,可通过BIOS层深度优化设置,实现计算密度提升:

硬盘控制器优化: 工作模式修改为AHCI模式,并允许存储驱动器支持高级串行ATA。

CPU模式优化: 禁用C-State节能状态,确保CPU快速响应计算请求,启用Turbo Boost 3.0技术,核心频率提升12%,

同时关闭节能模式,将能耗特性设为 Performance,提供最佳能源效率。

内存子系统调优: 配置NUMA节点本地化策略,降低跨Die访问延迟。

散热策略改进: 采用动态PID风扇控制算法,在80%负载下保持节点温度 $\leq 65^{\circ}\text{C}$,保证系统有效散热。

3.2 I/O 存储性能优化

HPC集群数据存储I/O性能除了依赖IB(InfiniBand)网络、存储硬件性能,还与存储架构有关,可通过构建三层存储体系实现数据全生命周期管理,具体包含:

热数据层,使用NVMe SSD如(Intel Optane),用于数据实时处理热交互。

温数据层,使用SAS HDD,采用并行文件系统,存储非实时数据。

冷数据层,使用磁带库,用于数据冷备份。

关键优化措施包括：

文件系统优化：文件系统关键需要维护好Home主目录，需做好NFS服务节点和Login登录节点的可用性维护^[20]。另外，可通过减少硬盘分区，让I/O负载尽可能分布到各个硬盘，对于与典型的文件不同的有特点的文件集合（如结构化文件、非结构化文件、大小、数量），创建一个针对其优化的文件系统。

I/O传输策略优化：采用Linux异步非阻塞I/O技术，允许进程启动多个I/O任务，读取请求后立即返回，而不必阻塞或等待任何操作完成。应用程序在后台读取到操作完成指令时执行其他处理，读取到响应到达则会生成线程回调信号号完成I/O操作。

数据访问模式优化：实施2MB大页内存管理，减少TLB缺失率，通过Stripe Size动态调整（1MB-4MB）匹配不同业务的数据访问模式。

硬件底层配置策略优化：以NVMe SSD为例，PCIe通道带宽理论模型可定义为：

$$BPCIe = \text{通道数} \times \text{单通道速率} \times \text{编码效率}$$

通道配置方面，在不升级硬件的情况下，可通过多盘绑定（如 RAID 0）确保 CPU /主板通道数充足，避免通道瓶颈（如 x8 通道下双盘并行理论带宽翻倍）。还可设置PCIe电源管理合理超时阈值（如 1ms），避免频繁状态切换引入延迟，减少空闲功耗。NVMe固件和控制器优化方面，设置电梯调度（SSD 内部队列排序）或多队列负载均衡，最小化逻辑块地址（LBA）跳转，降低访问延迟。设置控制器错误恢复机制，启用端到端数据保护（E2E CRC），但需平衡开销（增加约 5% 带宽消耗），适用于气象实时业务等高可靠性场景。

3.3 GPU 访存性能优化

访存优化的核心在于减少数据搬运和数据访问延时。

（1）优化数据搬运的主要手段是分块思想，以英伟达的GPU为例，访问内存为三级结构，分别是全局内存、共享内存和寄存器，分块方式打破传统计算机多级存储需要分层搬运数据的模式，而是尽可能将数据放在更近的存储结构多次利用，保证数据连续访问效率，减少浪费性的数据搬运过程，通过提升数据复用性提高存储性能。

（2）数据访问延时主要是由bank冲突造成的。为提升

内存宽带同时访问的效率，共享存储器被划分为多个大小相等的存储器模块，被称为bank。但在实际使用中，当一个线程组（warp）中不同线程同时访问一个bank地址就会发生冲突，冲突的发生和很多因素有关，其中数据类型、访问步长均可通过代码优化完成。从硬件层面主要从访问方式优化，可通过打开多播机制（multicast）方式解决冲突。如图2所示：

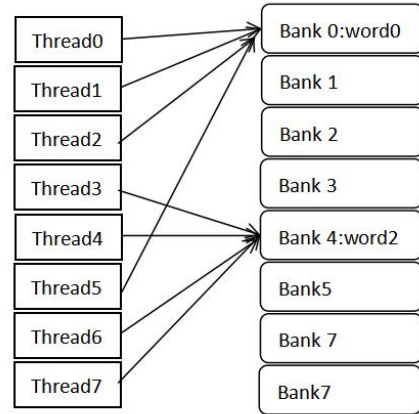


图2 多播机制解决多线程访问冲突情况图

正常情况下，多个线程同时访问bank0、bank4，会发生冲突。在多播机制（multicast）下，当一个warp中几个线程同时访问一个bank相同字地址时，会实时将该字地址广播给所有访问线程，可有效解决冲突延时。

（3）不同GPU架构优化方法。以NVIDIA分别于2017年、2020年发布的Volta、Ampere两种架构为例，二者在显存带宽、缓存容量、计算单元指令集、CUDA核心数量等都有一定差异。二者通用优化方法包含合并访存、数据重用、缓存优化、减少冗余访问等策略。针对不同架构的优化策略，针对Volta架构，可使用共享内存Bank冲突规避，寄存器分配优化（如通过循环展开和分块减少寄存器压力）等策略；针对Ampere 的优化，使用cudaMemcpyAsync将数据从全局内存异步复制到共享内存，与计算重叠，提高效率。启用Tensor Core 的结构化稀疏计算模式，使用 FP16/FP32 混合精度计算、HBM2e 硬件压缩减少显存占用，减少实际传输数据量。二者架构下的策略对比详见表1。

表1 NVIDIA Volta、Ampere 不同架构优化策略对比

优化点	Volta	Ampere
数据重用	依赖共享内存和寄存器	结合 L2 缓存和共享内存，减少全局访问
Bank 冲突	32 个 Bank，需严格对齐访问模式	64 个 Bank，冲突概率降低，支持更灵活访问
异步操作	无硬件支持，需软件实现重叠	硬件支持异步复制，减少 CPU-GPU 同步开销
稀疏计算	无原生支持，需软件实现	第三代 Tensor Core 原生支持结构化稀疏
混合精度	手动管理精度转换	自动混合精度（AMP）与 TF32 硬件加速
Bank 冲突	32 个 Bank，需严格对齐访问模式	64 个 Bank，冲突概率降低，支持更灵活访问

3.4 网络优化

网络通信效率在硬件层面极大依赖于设备的网络类型，

不同网络带宽和延时指标将直接影响通信传输能力。HPC数据传输优先选用QDR Infiniband、FDR Infiniband或RoCE v2等超高速网络类型,可使用DRB拥塞控制机制代替传统消息节流,提升网络延迟和吞吐量,实现存储节点与计算节点的零拷贝数据传输^[21],本文主要从拓扑结构优化进行分析。

在高性能计算机IB网络设计中主要依赖于设备选型,通常思路是采购单一大端口模块化交换机,该交换机具备数量

较多端口,直接连接所有计算节点、管理节点、I/O节点、存储节点服务器,优势是成本较低,缺点是模块交换机如故障则会影响整体系统运行。为提升通信效率以及可靠性,可将拓扑优化为小交换机堆叠胖数拓扑接口,通过多套小型交换机,采取堆叠树形架构,分为核心层、边缘层和应用层,优化后的结构架构见图3。

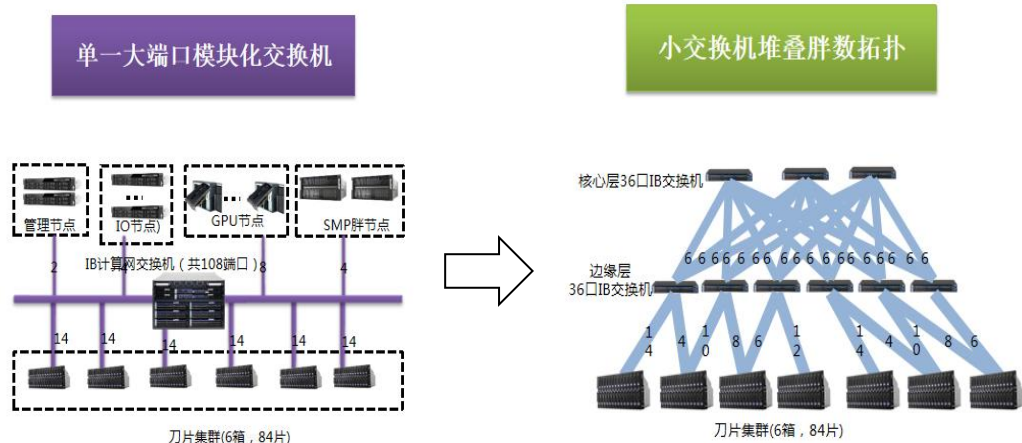


图3 省级气象高性能计算集群系统拓扑优化

4 动态作业调度优化

高效的资源调度是提升高性能计算集群吞吐量的核心,作业调度具备整合集群松散结构、异构软硬件资源,以及多用户提交任务统一调度的功能,集群作业调度系统可包含资源管理器、队列管理器、调度器,其核心为消息传递并行程序,目前业界包含MPI并行程序、PVM并行程序、任务级并行程序等,本文以气象HPC所用PBS调度系统为例,结合气象任务作业特性动态调整相关策略,从Linux操作系统配置优化、调度算法选择、动态资源分配以及负载均衡几个方面进行优化分析。

4.1 Linux 操作系统配置优化

内核参数调优,包含设置大页面,在内核编译时添加CONFIG_HUGETLB_PAGE 和CONFIG_HUGETLBFS 选项。开启网络优化,增大TCP缓冲区(net.core.rmem_max=16777216)、启用巨帧(MTU=9000)等。优化文件系统,使用XFS+noatime挂载选项减少元数据

开销,调整vm.dirty_ratio控制写回策略。

优化系统服务配置,在服务列表关闭多余的不用服务,如arpwatch、atd、autofs、chargen等。

开启进程与资源隔离策略,通过taskset或cpuset绑定进程到指定核心,避免跨核迁移,以提升CPU亲和性。优化系统交换分区配置,使用“Swapoff-a”命令,关闭swap分区项。开启实时性保障,对延迟敏感任务使用SCHED_FIFO调度策略(需CAP_SYS_NICE权限)。

4.2 智能调度算法策略

优化数据存储路径,预估作业任务量和计算节点性能需求,通过智能分配“短任务+空闲计算节点”组合方式,减少作业完成时间。同时采取资源回填(Backfill)策略,在Slurm/PBS中启用回填调度,填充空闲资源的小任务,提升资源利用率。

结合各用户业务重要性、实时性等特点,引用QoS(Quality of Service,服务质量)技术DiffServ(区分服务)模型如表2:

表2 三级作业调度模型

优先级	任务类型	资源保障策略
P0	同化模式预报、防灾减灾预警决策等重要实时业务	资源预留+抢占式调度
P1	气候模拟演算、多源数据融合等业务	公平份额+弹性配额,在队列优先策略的基础上,按用户组合理分配资源配额(如采用Hadoop公平调度器),避免资源垄断和长期性任务因资源枯竭而“饿死”
P2	零散型、短期性的科研类作业	根据集群作业规律,利用空闲资源进行运行,如Backfill策略

资源集合：设集群总资源为 $R=\{r_1, r_2, \dots, r_m\}$ ，例如 CPU、内存、GPU 等，总容量 $C=(C_1, C_2, \dots, C_m)$ 。

作业集合：作业分为三类优先级 $Q=\{Q_0, Q_1, Q_2\}$ ，每个作业 $J_i \in Q_k$ 的资源需求为 $D_i=(d_{i1}, d_{i2}, \dots, d_{im})$ 。

时间模型：离散时间片 $t \in T$ ，调度周期为 Δt 。

各优先级调度策略的形式化定义如下：

(1) P0级（资源预留+抢占式调度）。资源预留：为 Q_0 预留固定资源 $R_0 \subseteq R$ ，满足 $\sum_{j \in R_0} C_j \geq \max_{J_i \in Q_0} D_i$ 。

抢占条件：当 Q_0 作业到达且资源不足时，抢占 Q_1/Q_2 的资源，优先级满足 $Q_0 > Q_1 > Q_2$ 。

形式化约束：

$\forall t, \sum_{J_i \in Q_0(t)} D_i \leq R_0$ ，且 $R_0 \cap (R_1(t) \cup R_2(t)) = \emptyset$ ，其中 $R_k(t)$ 为 Q_k 在时刻 t 占用的资源。

(2) P1级（公平份额+弹性配额）。公平份额：为每

个用户/租户分配权重 ω_u ，资源配额 $S_u = \frac{\omega_u}{\sum \omega_u} \cdot R_1$ ，其中

$R_1 = R/R_0$ 。

弹性配额：允许临时借用空闲资源 R_i ，但上限为 $S_u + \alpha_u \cdot R_i$ ，弹性系数 $\alpha_u \in [0, 1]$ 。

形式化约束：

$$\forall u, \sum_{J_i \in Q_1^u(t)} D_i \leq \min(S_u + \alpha_u \cdot R_i(t), C_{\text{total}} - R_0) \quad (1)$$

(3) P2级（空闲资源运行）。资源分配：仅使用未被 Q_0/Q_1 占用的资源 $R_2(t) = R \setminus (R_0 \cup R_1(t))$ 。

作业调度：允许抢占 Q_2 内部作业，但不允许被 Q_1 抢占。

形式化约束：

$$\forall t, \sum_{J_i \in Q_2(t)} D_i \leq R_2(t) \quad (2)$$

三级调度优化模型参数选择及优化边界如表3所示：

表3 调度模型参数选择与优化边界

参数	理论依据	优化边界
预留资源 R_0	基于历史作业峰值 $\Pr(P_0 \leq R_0) \geq 99\%$ （极端值理论）	$R_0 \leq \beta \cdot C$ ，其中 $\beta \in [0.2, 0.5]$ ，避免资源浪费
抢占延迟 TP	上下文切换时间+检查点恢复时间	$TP \leq SLAP_0 - \text{JobRuntime}$
弹性系数 α_u	租户间博弈均衡	$\sum \alpha_u \leq 1$ ，防止资源过载
空闲资源探测周期 Δt	系统吞吐量与探测开销的权衡（排队论模型 M/M/c）	$\Delta t \propto \sqrt{\text{JobArrivalRate}}$

4.3 动态资源分配与负载均衡

通过开源容器编排引擎Kubernetes自动扩展计算节点，设置弹性伸缩，适应突发负载。对于高优先级紧急任务可采取抢占式调度保留资源，必要时终止低优先级任务。资源分配需考虑容错和负载均衡，在上述的气候演算等长期任务中，设置检查点策略，需要定期自动保存中间状态，还能根据当前节点队列负载情况，自动扩容或压缩算力资源，支持故障恢复和资源实施迁移，保障作业完成可靠性。同时设置资源队列组，对用户组资源配额设置上限值，防止单一用户组过度占用资源。

对于异构资源适配采取不同的调度手段，如区分不同作业任务对GPU、CPU的使用需求差异，弹性进行协同调度，对如雷达回波分析的图形类任务，需要GPU加速处理，可通过SLURM的--gres=gpu参数指定GPU资源，并设置独立队列管理15（SLURM在气象集群中的实际应用）。另外，一些如处理大量气象数据同化的业务，对I/O需求较大，可分配至配备NVMe SSD或并行文件系统的存储节点，避免与其他计算任务竞争带宽。

5 编译层级优化

编译优化直接影响代码执行效率，其优化需要结合集群硬件特性和用户作业程序进行调整。关键优化技术可包含混合编译优化、编译器配置优化、用户代码优化等方面。

5.1 气象编译器类别

编译器的选择主要依赖于硬件适配和用户业务特性，以

气象业务为例，分析编译器选择以下类别：

通用开源编译器选择如GCC、Clang/LLVM。根据CPU架构选择，如果集群使用x86架构，编译器可选择Intel编译器，如icc、icpc、ifort。针对气象业务各类雷达、卫星、云图等图形处理密集型业务需求，可选择GPU专用编译器，如专为NVIDIA GPU设计的CUDA编译器（NVCC）。

针对编程语言的编译器选择，针对气象应用较为广泛的Fortran编程语言，选择高性能Fortran编译器，如GNU编译器（gcc gfortran）、Intel Fortran（IFORT）；PGI编译器，如pgcc pgf90/95。

5.2 混合编译优化

在实际应用中，上述的编译器需组合使用提升编译效率，制定科学高效的混合编译策略是编译优化的核心。

根据计算特性，对计算密集型模块（如气象多模式运算）使用Intel编译器，对数据预处理和后处理需求量的可以组合使用开源编译器GCC、Clang。

区分异构资源使用需求，要充分考虑多线程编程OpenMP、分布式计算MPI和针对GPU并行编程的CUDA等编译器之间的兼容性，该因素直接影响多核与异构计算效率。在具体到某个气象业务作业上，要充分根据业务对计算、I/O性能、图形处理等需求，对上述三种并行编程技术合理地搭配组合使用，有效利用异构硬件优势，提升编程效率。例如某个既有分布式计算，又有大量气象数据图形处理需求的业务，可通过在MPI集群的每个节点上安装CUDA支持的GPU，来实现分布式计算与GPU加速的结合。使用支持GPU的MPI

库（如OpenMPI）来在多个节点上分配任务，并在每个节点上利用CUDA进行GPU加速，提升业务并行编程效率。

编译优化还需有效利用某些编译器动态调优、容错检查等功能，如使用AI驱动的编译器实现自动迭代编译参数，使用编译器检查点功能，保障长期性的任务运行可靠性。

5.3 编译器配置优化

编译器应根据集群硬件情况以及业务需求进行针对性的优化配置，本文主要以气象业务常用的GNU编译器为例，对优化选项进行分析。

编译优化需掌握适度原则，过度优化参数会降低运行速度，激进的优化策略会降低程序数值精度，导致计算结果误差率较大。通常GNU编译器优化级别包含O1、O2、O3三级，结合气象业务运算特点，通常选择O2优化级别。

5.3.1 配置选项优化

（1）内存与缓存优化。floop-block：分块优化循环以提高缓存利用率。Os，针对减小执行文件体积优化。floop-interchange：交换循环顺序优化内存访问模式。funroll-loops，循环展开优化。floop-unroll-and-jam：循环展开与合并（提升指令级并行）。

（2）架构适配。Intel CPU：march=skylake-avx512（适配Skylake架构）。mtune=haswell（针对Haswell微架构调优）。

AMD CPU：march=znver3（适配Zen3架构）。mprefer-vector-width=256（避免AVX-512降频问题）。

（3）链接时优化（LTO）。flto：启用链接时全局优化。

（4）指令集优化。march=native、-msse3。

5.3.2 核心计算优化

（1）浮点运算优化。Ofast：激进优化（包含-O3 -ffast-math，但可能牺牲精度）。

O3：最高级别通用优化（推荐基准选项）。ffast-math：放宽IEEE浮点规范约束，加速计算（需验证结果稳定性）。mfpmath=sse：强制使用SSE指令代替x87浮点单元（提升精度和性能）。

（2）向量化加速。march=native：自动适配本地CPU支持的指令集（如AVX2/AVX-512）。

msse4.2/-mavx2：显式启用SSE4.2/AVX2指令集（强制

生成SIMD代码）。fprefetch-loop-arrays：预取数组数据到缓存（适合大内存循环）。

（3）并行计算支持。fopenmp：启用OpenMP多线程并行（需代码支持OpenMP指令）。

floop-parallelize-all：自动并行化循环（需结合-free-parallelize-loops）。

6 代码级优化

编程中首先要保障编写代码的规范性，对于不规范的代码，过高的优化级别反而会导致程序运行错误。代码可优化的内容具体包含循环优化、移位乘除、循环不变计算、函数优化、条件判断句式优化等。代码优化主要针对具体的某个气象业务算法程序深入研究，本文仅从以下常规代码优化分析。

一是全局变量的调用。由于编译器不能将程序中的全局变量存储在寄存器，因此调用全局变量需要额外读取和存储，因此在重要循环中建议尽量减少全局变量的使用，或者将全局变量值转为局部变量。

二是循环语句的合理优化，保证长循环在内，短循环在外，用switch()函数替代if...else...函数，通常带参数的宏定义效率比函数更高，因此简单的运算可以用宏定义来完成。

三是灵活选用合适的数据结构，如插入、删除动作较多，优先选用链表。因大部分编译器使用指针比数组执行效率更高，因此可用指针运算代替数组索引，较少存储占用空间，生成代码更加简洁。如使用数组，减少频繁函数调用，将递归调用改为内循环，查询方式可采用二分法查询方式，提升查询效率。

7 省级气象 HPC 优化应用情况

按照上述多维度全栈式策略对某省级气象HPC进行优化，最终HPC算力性能实现有效提升，其中内存平均访问速率提升1倍、作业调度平均等待（Pending）时间缩短1.5倍，调度时间、频次也得到有效优化，业务作业平均完成时间缩短32%、突发任务资源响应时间≤30秒，优化后系统运行监控如图4—6所示。

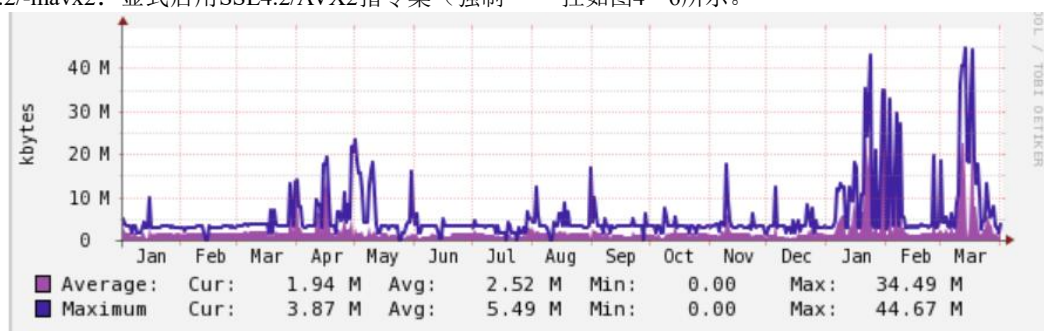


图4 某用户组的内存访问速率提升情况



图 5 集群作业队列等待 (Pending) 时间、拥堵频次变化情况

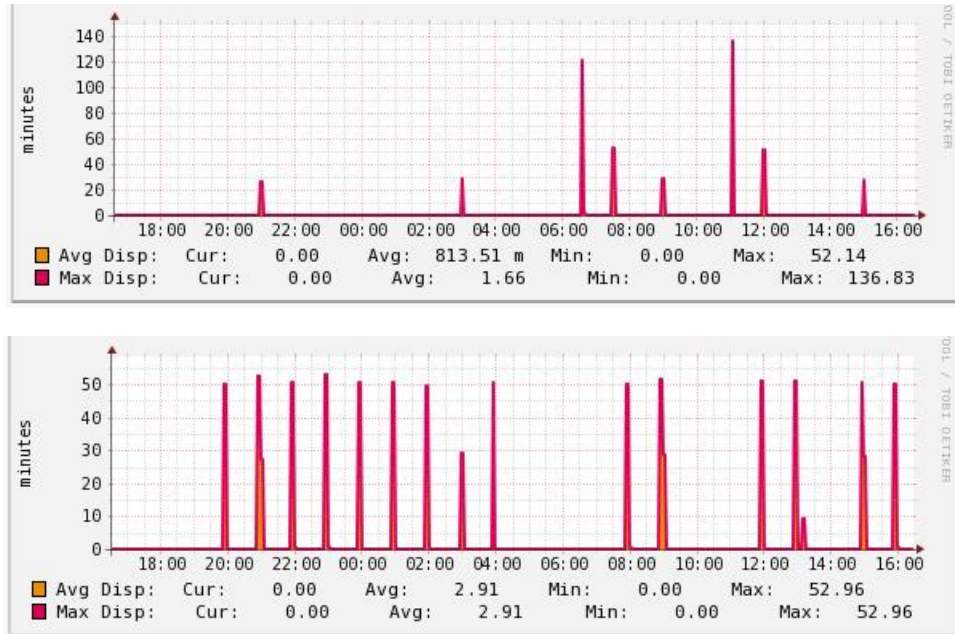


图 6 集群队列调度时间、频次变化情况

8 结语

本文基于某省级气象HPC以及气象业务特点,从多种维度优化策略进行研究,通过多维协同策略改进,集群算力提升效益明显,总结HPC集群优化需遵循层次分析、链路协同的总体原则,在硬件层根据业务架构,在保障负载均衡的前提下最大化异构资源利用率,操作系统需通过内核调优和资源隔离提升质效,调度层面满足动态匹配需求和资源饱满供给能力,编译方面向目标硬件生成高效机器执行码,代码层

面要针对具体的业务程序,保证代码规范编写,提升读写效率。随着人工智能、物联网等技术飞速发展,HPC行业也逐步开展动态调度优化策略机器学习、工具链异构存储资源全栈性能分析、性能功耗动态优化等技术应用研究,本文对气象部门以及相关行业的高性能计算机应用有一定借鉴作用。

参考文献

- [1] 王洁.高性能计算应更注重与 AI、大数据等新兴技术深度融合

- [N],中国信息化周报,2024-10-21.
- [2] 张云泉,袁良,袁国兴,等.2024 年中国高性能计算机发展现状分析[J].计算机工程与科学,2024,46(12):2091-2098.
- [3] 高岳,郭春华,米雪,等.国产 HPC 与 AI 芯片制造装备技术现状与发展策略分析[J].电子工业专用设备,2025,54(01):1-6+27.
- [4] 杨敏,何芸,许涛,等.高性能 GPU 计算集群应用体系建设[J].信息系统工程,2025(03):102-105.
- [5] 王亚娟.高性能计算平台在气象灾害预警系统中的应用[J].中国科技信息,2025(06):82-84.
- [6] 张宇,陈德辉.谈谈高性能计算机对数值天气预报发展的重要技术支撑作用[J].科研信息化技术与应用,2010,1(04):9-19+8.
- [7] 明宇.高性能异构计算的任务调度技术[D].成都:电子科技大学,2024.
- [8] 张志坚,伍光胜,孙伟忠,等.IBM Flex P460 高性能计算机系统及气象应用[J].现代计算机,2016,(09):51-55.
- [9] 徐顺,王武,张鉴,等.面向异构计算的高性能计算算法与软件[J].软件学报,2021 ,32 (08):2365-2376.
- [10] 杨昕,沈文海.“东数西算”趋势下的气象算力网络和算力服务体系架构[J].数据与计算发展前沿,2022 ,4 (05):50-59.
- [11] 贾晨刚,蔡玉婷,黄勤,等.陕西气象高性能计算系统扩容设计[J].微型电脑应用,2025 ,41 (02):309-312.
- [12] 沈瑜,李娟,常飏,等.高性能计算机统一资源管理系统的设计与实现[J].计算机技术与自动化,2014,33(01):83-90.
- [13] 顾文静,常飏,李娟,等.高性能计算机资源管理系统改进设计与实现[J].计算机技术与自动化,2017,36(04):104-109.
- [14] 黄小猛,林岩鑫,熊巍,等.数值预报 AI 气象大模型国际发展动态研究[J].大气科学学报,2024 ,47 (01):46-54.
- [15] 雷鸣,宫志宏,郭阳,等.气象人工智能通用算法库的设计与实现[J].微型电脑应用,2024 ,40 (05):1-4.
- [16] 李江南.基于高性能计算集群的蚁群算法改进研究[D].南昌:南昌大学,2018:45-48.
- [17] 吴璨,肖海力,王小宁,等.面向高性能计算环境的智能任务编排架构研究[J].数据与计算发展前沿(中英文),2025 ,7 (01):99-107.
- [18] 田彦明,候永生,高超.高性能并行计算中节点负载均衡分配算法[J].信息技术与信息化,2024 (11):40-43.
- [19] 吕小敏,刘钊,蔡蕙伊,等.面向国产神威众核架构的区域气候模式 CWRF 性能优化技术[J].电子技术应用,2022,48(01):31-38.
- [20] 沈瑜,孙婧,李娟.中国气象局高性能计算机系统高可靠性设计[J].信息安全与技术,2013,4(06):42-45+57.
- [21] 张众钦.基于 InfiniBand 的高性能计算机技术研究[J].信息记录材料,2021,22(11):231-232.

作者简介

王晓东（1988-），男，学士，工程师，研究方向：气象信息技术研究。

Research on the computing power optimization method of meteorological HPC based on multi-dimensional cooperative strategy

WANG Xiaodong, DENG Weihua, LI Hongkang, XIONG Honghua

Jiangxi Provincial Meteorological Data Center, Nanchang, JiangXi 330096, China

Abstract: With the rapid development of machine learning and big data technology, the demand for computing resources in the field of meteorology is increasing exponentially. Under the condition of limited hardware resources, how to improve the computing power utilization efficiency of high-performance computing cluster (HPC) has become a key research topic. Based on the architecture and business characteristics of a provincial meteorological HPC cluster, this paper puts forward a multi-dimensional collaborative optimization framework, and conducts research from four core levels: hardware resource scheduling, system configuration and tuning, dynamic task management and compilation efficiency improvement. By building a hierarchical optimization model, the performance of storage I / O throughput, job scheduling efficiency, and compilation and execution efficiency are significantly improved. The research results provide a reusable technical path for the efficiency optimization of HPC system in

the meteorological field.

Keywords: High performance computing, Computing power optimization, Meteorological application, Heterogeneous resource scheduling